

Description

METHOD FOR DETECTING EXTENDED RANGE MOTION AND COUNTING MOVING OBJECTS USING AN ACOUSTICS MICROPHONE ARRAY

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit under 35 USC 119(e) of provisional application 60/319,652 filed October 29, 2002 and of provisional application 60/320,214 filed May 23, 2003, the entire file wrapper contents of which provisional applications are herein incorporated by reference as though fully set forth at length.

FEDERAL RESEARCH STATEMENT

[0002] The invention described herein may be made, used, or licensed by or for the United States Government for government purposes without the payment of any royalties thereon or therefore.

BACKGROUND OF INVENTION

[0003] *FIELD OF THE INVENTION*

[0004] This invention generally relates to detection of moving targets, and more specifically to detection of moving targets using one or more acoustic arrays. More particularly, this invention pertains to a system that uses conventional

and/or adaptive beamforming to locate and count multiple targets in dynamic, noisy, and highly mobile environments.

[0005] *BACKGROUND OF THE INVENTION*

[0006] Acoustic over watch sensor systems are used to provide extended range surveillance, detection and identification for force protection and tactical security on the battlefield. A network of over watch sensors remotely deployed in conjunction with a central processing node (or gateway) can provide early warning and assessment of enemy threats, near real-time situational awareness to commanders, and may reduce potential hazards to the soldier. When a network of acoustic sensors is used, measurements from distributed arrays are reported to a "fusion" center that sorts the sensor measurements and estimates the location and kinematics of identified sources of interest.

[0007] An essential requirement for unattended acoustic sensors systems used for battlefield surveillance is the capability to locate and track multiple targets in dynamic, noisy, and highly mobile environments. Today's state-of-the-art acoustic sensor systems report measurements from diverse sources comprising targets of interest, background noise (or clutter), and noise due to local environmental conditions. However, complex data association logic is required to sort out sensor measurements into general categories such as targets of interest, sources that are not of interest and false detections that cannot be correlated over time.

[0008] Target tracking in noisy, cluttered environments is treated as a problem in associating detected target observations with target tracks. These

observations could be comprised of false target detections or an unknown number of actual target detections. Consequently, the performance of these methods is fundamentally dependent on the system's target detection capability.

[0009] Conventional and adaptive beamforming algorithms are commonly used to enhance the directivity and direction finding capability of a single sensor array and are essential in realizing the performance benchmarks required in detecting and localizing remote targets of interest. These spatial filtering methods allow position-related measurements such as bearing to a target to be computed from the steered response of an acoustic array. One of the most challenging areas of the multi-sensor performance problem is the beam-pointing control of an acoustic array in dense multiple target environments, where closely spaced targets, highly maneuvering targets, and targets in the presence of multi-path should be resolved. Factors that contribute to this challenge are finite sensor resolution, inconsistent beam shape, false alarms, complex target formations, and the mobile nature of target signatures.

[0010] The use of multiple acoustic sensors in array configurations for the purpose of forming specific narrow look directions (beamforming) and using such means to extend the detection range of individual microphones has been implemented and tested for land acoustic vehicle detection. Such implementations have been successful in identifying the presence of multiple targets in the local vicinity of the acoustic array by measuring multiple bearing directions to each target with respect to the acoustic array. Some implementations also determine a passive range estimation to target. By

combining this passive range estimation with bearing to target information, multiple tracks to targets can be maintained. The number of targets that can be effectively tracked by an acoustic array using adaptive beamforming methods is widely recognized as the quantity $(N - 1)$ where N is the number of microphones in the array. Consequently, if more than $N-1$ targets enter the surveillance area, confusion potentially exists and the ability to individually track each vehicle cannot be maintained.

[0011] In practice, a basic problem of accountability occurs when a count for the number of targets that pass by the vicinity of the acoustic array is desired. Often a user must examine the ongoing real-time output of the sensor array to determine a best guess for the number of vehicles that were tracked during any particular timeframe. Therefore, a man-in-the-loop is required to perform very careful examination of the ongoing information. Often, the tracks are not maintained continuously and without high fidelity (depending on range from the array and spacing of targets). Consequently, it is difficult to know which tracks should be "counted".

[0012] If more than one acoustic array is required to cover a greater area, accountability of target count becomes very difficult. To mitigate this issue, multiple microphone arrays are generally networked together to improve the effective area of coverage. Networked acoustic arrays that combine their respective bearing information to determine more accurate target tracking (using triangulation) should agree on which targets they are tracking from remote locations. Deciding if the target is the same target being viewed from multiple remote locations is inherently difficult, often creating "ghost" targets, which are targets that don't exist.

[0013] The sharing of tracking information from multiple sensor sites requires that each site resolve individual targets. In scenarios where targets are closely spaced or moving abreast relative to a sensor node, the sensors do not "hear" the same target over portions of the track through the sensor field, therefore, tracking and target counting accountability is very difficult.

[0014] What is needed are signal processing techniques that would provide over watch sentries with the ability to autonomously and remotely count the number of vehicles passing through a predetermined field of view. The need for such a system has heretofore remained unsatisfied.

SUMMARY OF INVENTION

[0015] The present invention satisfies this need, and presents a method for detecting extended range motion and counting moving objects (sound sources) using an acoustics microphone array. The acoustic microphone array employs an optimized beamforming technique to create specially focused listening directions that function as a "motion detector" and "trip line counter". Algorithms detect and confirm the presence of objects moving through the beams and perform other algorithmic tests to verify that the object is a valid object and should be counted. In addition, algorithms verify valid directional information of the key target features within the acoustic beams.

[0016] The present method represents a new approach to the problem of counting targets using one or more acoustic arrays and can effectively count targets without human intervention. The present method allows the addition of an "acoustic trip line counter " around a surveillance area as a "building block"

for each acoustic node placed in the area. The user implements the present method by selecting a look direction across which the "motion detection" and resulting counting may occur. Two or more acoustic arrays that look down the same avenue can extend the range of the "acoustic trip line" to cover large battlefield frontage. Measurements are partitioned into sets of observations and are confirmed so that background noise and other false detections are rejected while estimates for the number of targets can be computed. While the present invention is discussed in terms of vehicles in a battlefield situation, the present invention can be applied to moving objects in any scenario.

[0017] In many projected remote sensor deployment concepts, there exists a need to form a perimeter around the sensor surveillance area. Implementation of the present method allows the formation of an effective surveillance perimeter using multiple acoustic sensor nodes. The information from the present method is useful as a powerful stand-alone technique to count targets in specific surveillance zones. The present method may also be added to existing system level tracking solutions implemented for the purpose of better defining the location of multiple vehicles in confusing multiple vehicle scenarios.

[0018] The present method is also a solution for resolving track loss deficiencies inherent to tracking individual vehicles in cluttered multiple target environments within the proximity of the closest point of approach to an acoustic sensor. As an example, the present method may be used in conjunction with an array baseline comparable to that of the integrated acoustic sensor system (IAS) developed by Textron Systems Division.

[0019] The present method introduces a way of alleviating uncertainties regarding the number of vehicles being tracked at closest point of approach based solely on individual sensor information. The performance of the present method is ensured using signal processing techniques that provide the capability of both detecting the presence of every vehicle that approaches an acoustic array and obtaining a highly accurate vehicle count.

[0020] In dense target environments, the detected acoustic signals are contaminated by the engine noise from other vehicles that are often louder than the target of interest. The problem of separating signals from individual targets in multiple target scenarios requires a robust adaptive approach to suppress interference from other vehicles closely spaced to the target of interest. The present method solves this issue through an acoustic "trip line" algorithm that addresses problems with the array's limited resolution due to its finite aperture. The method also introduces noise suppression that excludes undesired signals and overcomes the inability to account for source motion.

[0021] An implementation of the "trip line" approach of the present method utilizes both conventional and adaptive beamforming algorithms that form a refined spatial filter to capture the harmonic content (key features) of individual targets that pass through a fixed line in bearing. The power from narrowband signal components in both the conventionally and adaptively formed beams is monitored so that a target crossing a line of bearing through one of the designated beams triggers the event of a target count. Conventional beamforming is used to "coarsely" filter and estimate the direction of arrival of narrowband acoustic sources impinging on the circular array of

microphones. The "fixed" beam pointing and highly selective spatial filtering is achieved using a minimum variance distortionless response (MVDR) beamforming algorithm in which the spatial filtering function is frequently updated in response to changes in the local environment. This adaptive weight computation ensures a beamformer response that approaches the theoretically optimum solution by reducing (and often canceling) the effects of interfering signals. Significant performance advantages are gained using the adaptive beamforming algorithm. The adaptive strategy also accounts for movements in source direction by frequently updating the spatial characteristics of the array to obtain an optimal signal to noise gain in the direction of the target.

BRIEF DESCRIPTION OF DRAWINGS

[0022] The various features of the present invention and the manner of attaining them will be described in greater detail with reference to the following description, claims, and drawings, wherein reference numerals are reused, where appropriate, to indicate a correspondence between the referenced items, and wherein:

[0023] FIG. 1 is a schematic illustration of an exemplary operating environment in which a moving target counting method of the present invention can be used;

[0024] FIG. 2 is a graph illustrating the use of a trip line by the moving target counting system of FIG. 1;

[0025] FIG. 3 is a process flow chart illustrating a method of operation of the moving

target counting algorithm of FIG. 1;

[0026] FIG. 4 is a process flow chart illustrating a method of operation of the conventional beamformer module of the moving target counting algorithm;

[0027] FIG. 5 is a process flow chart illustrating a method of operation of the adaptive beamformer module of the moving target counting algorithm; and

[0028] FIGS. 6A, 6B, 6C, 6D, 6E, 6F, 6G, and 6H are a series of graphs illustrating the operation of the moving target counting method in an exemplary deployment.

DETAILED DESCRIPTION

[0029] FIG. 1 illustrates an exemplary high-level architecture of an acoustic sensor system comprising a moving target counting system 10 that utilizes an algorithm for detecting extended range motion and counting moving objects using an acoustics microphone array. System 10 includes a software programming code or computer program product that is typically embedded within, or installed on an Analog to Digital Converter (ADC) / Processor memory module 15. The microphone array 20 comprises a set of microphones such as microphone 1- (#25), microphone 2- (#30), microphone 3- (# 35), microphone 4- (#40), up to and including microphone n- (# 45). The microphone array 20 provides acoustic information to system 10 for detecting, tracking, counting and classifying targets. The compass 50 and GPS module 55 also provide information to the system 10 for location and steering purposes.

[0030] An acoustic sensor system can be used to perform direction finding using a

fixed array strategy to optimize the filtering of microphone array 20 in a steered focus direction.

[0031] Beamforming generally used to steer microphone array 20 in both azimuth and elevation and measure the resulting output power from a finite set of sampled data. The spatial filtering operation passes signals from fixed directions of interest and rejects noise from all other directions. In frequency domain beamforming, the response of microphone array 20 is steered by forming a linear combination of the sensor outputs

[0032]

$$y(t) = \sum_{m=1}^M w_m^* x_m(t) = \mathbf{w}^H \mathbf{x}(t) \quad (1)$$

[0033] where $\mathbf{w} \in \mathbb{C}^n$ is the weight vector of spatial filter coefficients and \mathbf{x} is the vector obtained from the Short Time Fourier Transform (STFT) of the sampled signal $x(t)$ computed at discrete frequency bins.

[0034] Given a set of samples $\mathbf{x}(t)$ observed over a time interval T , the output power is measured as:

[0035]

$$\mathbf{P}(\mathbf{w}) = \frac{1}{T} \sum_{t=1}^T |y(t)|^2 = \frac{1}{T} \sum_{t=1}^T \mathbf{w}^H \mathbf{x}(t) \mathbf{x}^H(t) \mathbf{w} = \mathbf{w}^H \mathbf{R} \mathbf{w} \quad (2)$$

[0036] where \mathbf{R} is the signal spatial covariance matrix $\mathbf{R} = E\{\mathbf{x}(t) \mathbf{x}^H(t)\}$ and $E\{\cdot\}$ denotes the statistical expectation. In practice, the signal covariance matrix is not known and should be estimated from the spectral data:

[0037]

$$\bar{\mathbf{R}} = \frac{1}{T} \sum_{t=1}^T \mathbf{x}(t) \mathbf{x}^H(t). \quad (3)$$

[0038] A conventional beamformer (CBF) steers a single beam in the direction of a signal of interest and places nulls in fixed locations independent of interfering noise sources. Let $\mathbf{a}(\theta) \in \mathbb{C}^N$ denote the response of a microphone array 20 of arbitrary geometry to a plane wave of unit amplitude arriving from some focus direction θ . A measurement of the output of microphone array 20 $\mathbf{x}(t) \in \mathbb{C}^N$ to a narrowband source $s(t)$ in the far field of the microphone array 20 is given as:

$$[0039] \quad \mathbf{x}(t) = \mathbf{a}(\theta)s(t) + \mathbf{\zeta}(t). \quad (4)$$

[0040] The additive noise $\mathbf{\zeta}(t)$ represents undesired signals. For the conventional beamformer, the output described in (1) is given as

$$[0041] \quad y(t) = \mathbf{w}^H \mathbf{a}(\theta)s(t) + \mathbf{w}^H \mathbf{\zeta}(t). \quad (5)$$

[0042] The problem of maximizing the output power then becomes

$$[0043] \quad \max_{\mathbf{w}} E\{\mathbf{w}^H \mathbf{x}(t) \mathbf{x}^H(t) \mathbf{w}\} = \max_{\mathbf{w}} \mathbf{w}^H E\{\mathbf{x}(t) \mathbf{x}^H(t)\} \mathbf{w} = \max_{\mathbf{w}} \left\{ E|s(t)|^2 |\mathbf{w}^H \mathbf{a}(\theta)|^2 + \sigma^2 |\mathbf{w}|^2 \right\}. \quad (6)$$

[0044] The spatial coefficients are computed at discrete frequencies to ensure that the weighted signal components from waves propagating over the microphone array 20 may add constructively with unit magnitude and zero phase. This property is referred to as the distortionless response of a microphone array 20 and is satisfied when the gain of the weighed response is such that

[0045]
$$|\mathbf{w}^H \mathbf{a}(\varphi)| = 1. \quad (7)$$

[0046] In general, values for the weights that minimize the noise gain $\mathbf{w}^H \mathbf{E}\{\zeta(t)\zeta^H(t)\}\mathbf{w}$ are desired. For the conventional beamformer, the noise $\zeta(t)$ is assumed to be uncorrelated among microphones.

[0047] A non-trivial solution is obtained when the norm of the weight vector is constrained to unity when performing the maximization in Equation 6, resulting in an expression for the weight vector:

[0048]
$$\mathbf{w} = \frac{\mathbf{a}(\varphi)}{[\mathbf{a}^H(\varphi)\mathbf{a}(\varphi)]^{1/2}}. \quad (8)$$

[0049] The spatial spectrum is found by substituting Equation 8 into Equation 2:

[0050]
$$\mathbf{P}(\varphi) = \frac{\mathbf{a}^H(\varphi)\overline{\mathbf{R}}\mathbf{a}(\varphi)}{\mathbf{a}^H(\varphi)\mathbf{a}(\varphi)} \quad (9)$$

[0051] where $\overline{\mathbf{R}}$ is defined in Equation 3. The conventional beamformer

$\overline{\mathbf{R}}$

maximizes the output power in Equation 2 for a desired signal direction and places nulls in fixed locations independent of interfering noise sources.

[0052] In general, system 10 requires the ability to resolve the power of two acoustic sources spaced closer than the natural beam width of the microphone array 20. This is accomplished by implementing an MVDR beamformer that is capable of adapting to the noise levels found within

multiple target environments using microphone array 20. Using this approach, the weights are chosen to minimize the beamformer power in Equation 2 subject to the distortionless response criteria in Equation 7. Assuming that the beamformer response $\hat{a}(\#)$ and the spatial covariance matrix $R=E\{x(t)x^H(t)\}$ are known, the weights may be chosen as the optimal solution of

$$\begin{aligned} \text{[0053]} \quad & \text{minimize} \quad \mathbf{P}(\mathbf{w}) \\ & \text{such that} \quad \mathbf{w}^H \mathbf{a}(\varphi) = 1 \end{aligned} \quad (10)$$

[0054] where $P(w)$ is defined in Equation 2. The MVDR beamformer makes use of an estimated signal covariance matrix (defined in Equation 3) in

$$\bar{\mathbf{R}}$$

place of R for the output power in (10). The resulting beamformer is chosen as the solution to

$$\begin{aligned} \text{[0055]} \quad & \text{minimize} \quad \mathbf{w}^H \bar{\mathbf{R}} \mathbf{w} \\ & \text{such that} \quad \mathbf{w}^H \mathbf{a}(\varphi) = 1 \end{aligned} \quad (11)$$

[0056] The optimal value for the weight vector w can be found using a complex Lagrange multiplier λ in conjunction with the unconstrained cost function

$$\mathbf{J}(\mathbf{w}) = \mathbf{w}^H \bar{\mathbf{R}} \mathbf{w} + \text{Re} \left[\lambda \left(\mathbf{w}^H \hat{\mathbf{a}}(\varphi) - 1 \right) \right]$$

and minimizing over w and λ . The necessary conditions are found to be

$$\partial \mathbf{J} / \partial \mathbf{w} = \bar{\mathbf{R}} \mathbf{w} + \lambda \hat{\mathbf{a}}(\varphi) = 0 \quad \text{and} \quad \partial \mathbf{J} / \partial \lambda = \mathbf{w}^H \hat{\mathbf{a}}(\varphi) - 1 = 0$$

that when combined yield the analytical solution:

[0057]
$$\mathbf{w}_{\text{MVDR}} = \frac{\bar{\mathbf{R}}^{-1} \hat{\mathbf{a}}(\varphi)}{\hat{\mathbf{a}}^H(\varphi) \bar{\mathbf{R}}^{-1} \hat{\mathbf{a}}(\varphi)} \quad (12)$$

[0058] The MVDR spatial spectrum is found by inserting the expression for \mathbf{w}_{MVDR} into Equation 2:

[0059]
$$\mathbf{P}(\varphi)_{\text{MVDR}} = \frac{1}{\hat{\mathbf{a}}^H(\varphi) \bar{\mathbf{R}}^{-1} \hat{\mathbf{a}}(\varphi)} \quad (13)$$

[0060] The MVDR beamformer minimizes the power contribution of the noise and any signals coming from directions other than the direction of interest #, while maintaining a fixed gain in the desired focus direction. In this way, nulls are automatically placed in directions of interfering noise sources while a distortionless response is maintained in a prescribed look direction. Although the resolution of the beamformer is still dependent on the aperture of the microphone array 20 and signal to noise ratio (SNR) of both the source and interfering signal components, the spectral leakage from closely spaced sources is generally reduced.

[0061] The adaptive MVDR beamforming solution of system 10 (also referenced as adaptive beamforming or adaptive beamformer) measures the spatial covariance matrix rather than making assumptions about the noise. The matrix $\bar{\mathbf{R}}$ in system 10 varies due to the dynamic nature of the

$\bar{\mathbf{R}}$

desired signal components and to the noise typically found in a multiple

target environment. In general, the estimate for $\bar{\mathbf{R}}$ can be time

$$\bar{\mathbf{R}}$$

averaged as in Equation 3 to get a better estimate of the signal and to effectively reduce the noise. However, the time-varying nature of the data leads to an estimate for $\bar{\mathbf{R}}$ that is susceptible to errors when a small

$$\bar{\mathbf{R}}$$

number of samples are averaged in the computation.

[0062] When the number of averages is less than the number of elements in microphone array 20, the estimate for the spatial covariance matrix results in an ill conditioned (or singular) matrix. When this occurs, the weights are adjusted in such a way that the noise cancels the signal even when the look constraint is satisfied. Nevertheless, it is desirable to continuously update the spatial characteristics of the microphone array 20 using the least amount of averages to account for movement in source direction. The object is to maintain good noise suppression of interfering signals even if the signal of interest and interfering noise occupy the same temporal frequency band. This problem is inherent to multiple target tracking applications when similar vehicles are tracked over time.

[0063] The MVDR beamformer is commonly desensitized to estimation errors by injecting a small amount of white noise along the diagonal elements of $\bar{\mathbf{R}}$ to ensure its invertability. The process of adding a scaled identity

$$\bar{\mathbf{R}}$$

matrix to the covariance estimate is typically referred to as regularization or diagonal loading. The degradation in noise suppression that results for large μ^{-1} is resolved by adding a constraint on the noise response of the

$\bar{\mathbf{R}}$

beamformer in the optimization problem Equation 11 as:

$$\begin{aligned} \text{[0064]} \quad & \text{minimize} \quad \mathbf{w}^H \bar{\mathbf{R}} \mathbf{w} + \mu \mathbf{w}^H \mathbf{w} \\ & \text{such that} \quad \mathbf{w}^H \mathbf{a}(\varphi) = 1 \end{aligned} \quad (14)$$

[0065] where $\mu > 0$ is a fixed parameter used to detune the beamformer response by penalizing large values of \mathbf{w} . The solution is written as

$$\text{[0066]} \quad \mathbf{w}_{\text{MVDR_DL}} = \frac{(\bar{\mathbf{R}} + \mu \mathbf{I})^{-1} \mathbf{a}(\varphi)}{\mathbf{a}^H(\varphi) (\bar{\mathbf{R}} + \mu \mathbf{I})^{-1} \mathbf{a}(\varphi)} \quad (15)$$

[0067] System 10 uses a regularization parameter μ that is adjusted relative to the average background noise level computed at each sensor element. A rank order filter assigns the noise level around the frequency in question as the 30th percentile value of sorted spectral magnitudes at each frequency bin within a 25 Hz window. The actual value of μ is selected to be within 5dB of the background noise value that is averaged over the number of microphones in the microphone array 20.

[0068] The target counting algorithm of system 10 takes advantage of the capability of a microphone array 20 to detect the presence of vehicles by measuring the harmonic content in a pre-determined "look-direction" along a fixed line of bearing (as illustrated in FIG. 2). The "trip line" is formed with a narrow

eam pointing response that is derived from the conventional and adaptive response of the microphone array 20 to obtain automatic counts to individual targets passing through a fixed line in bearing. An adaptive beamformer steers deep nulls to suppress harmonic content in the directions of interferers while bearing estimation using a conventional fixed-beam method is used to compute the direction of arrival of narrowband sources that are passed from the predetermined "look-direction". The power in beams steered in the "trip line" direction is monitored such that targets crossing through the predetermined steering angle automatically trigger the event of a target count.

[0069] The adaptive "trip line" beam is pointed in the direction of a target's closest point of approach to the sensor since the range to the target is the shortest and the target is generally loudest at this point. In addition, the highest rate of change in bearing to the target occurs at this time. The acoustic sensors detect the presence of a target by measuring the harmonic content in the prescribed "look-direction". The auxiliary logic of system 10 is used to ensure a minimum number of false counts. The adaptive trip line is made possible through the implementation of an MVDR beamformer that allows for integrated main lobe and side lobe cancellation determined from the nature of the source signal(s). When a vehicle moves into the path of the steered beam, sound energy created by the passing vehicle passes undistorted in the "trip line" direction due to the enhanced directivity of the beamformer in this direction. The rationale for this response is derived directly from the definition of the MVDR beamformer (Equation 11) that states that the beamformer should pass signals undistorted from the direction in which the beam is pointed while minimizing the total beam power in all directions.

als that are not in the look direction may be automatically suppressed by the beamformer to minimize the total beam power.

[0070] Through the significant functional capability of system 10, an acoustic microphone array 20 can be positioned in a remote location off the expected path that multiple vehicles may traverse. The user can then select an intended look direction to make a count of vehicles as well as verify or classify the vehicle type. The functional improvement is provided by combining an acoustic feature extractor with detailed location information of the vehicle that together verify the presence of a vehicle in the specific acoustic beamforming "look direction". Multiple acoustic beams are formed using a beamforming technique referenced as Minimum Variance Distortionless Response (MVDR) or adaptive beamforming, which provides directed narrow acoustic look directions with the least amount of noise interference from surrounding sources not in the beamformer look direction. The use of multiple beams is intended to provide additional confidence of the continuity of the vehicle motion through the counting "look direction" and is in essence a motion detector.

[0071] The user places the acoustic microphone array 20, comprising n number of microphones and uses the algorithm of system 10 to electronically steer the beams into the desired look direction or path across which vehicles may be counted. The range of the system is proportional to the number of microphones in the acoustic microphone array 20. Many configurations of the acoustic microphone array 20 are possible to produce the beamformer response required for target counting. If the look direction is too wide, target counting may not be accurate. For example, the use of a 45 degree look

direction works well with a 4 foot, 5 microphone array 20. The ability to form well defined beams within the look direction forms the basis of an acoustic "motion detector".

[0072] Once the vehicle enters the first beam, system 10 is automatically triggered to search again in the center beam, and then once again as the vehicle exits the outer beam. If this criteria is met, the essence of target motion detection is satisfied.

[0073] A method 400 of operation of system 10 is illustrated by the process flowchart of FIG. 3. System 10 obtains sampled acoustic data from microphone array 20 at block 405. For each channel, system 10 computes the short time Fourier Transform (STFT) at block 410. At block 415, system 10 computes the power spectrum for each channel. The power spectrum computed at block 415 is used by both the conventional beamformer module 420 and the adaptive beamformer module 425. The method of the conventional beamformer module is illustrated in more detail in FIG. 4; the method of the adaptive beamformer module 425 is illustrated in more detail in FIG. 5.

[0074] The output of the adaptive beamformer module 425 is examined at block 430 and all spectral components in each beam that are greater than a predetermined threshold are retained. The conventional beamformer module 420 has two outputs. Output A 435 is the bearing to the single loudest spectral component. Output B 440 comprises the bearings to all spectral components. System 10 assumes actual bearings to peaks passed by the adaptive beamformer are along the steered look direction. System 10 then

correlates these peaks with components found using the conventional beamformer (output B 440) and associates peaks existing in both CBF beams and MVDR beams. At block 445, system 10 combines the bearings to all spectral components (output B 440) and the spectral components greater than a predetermined threshold (block 430) to assign bearings to spectral components.

[0075] System 10 retains spectral components within a bearing tolerance of each look direction and counts the total number of components in each direction at block 445. System 10 retains peaks in the MVDR beams that are $\pm\mu$ degrees with respect to the principal look direction. System 10 also retains peaks in the MVDR trip line if they are within a tolerance of $\pm\#^\circ$ with respect to the steered trip line look direction. This tolerance determines the narrowness of the resulting trip line.

[0076] A spectral component is registered as a moving target if the trip line is triggered. The trip line is triggered when the trip line beam contains the most number of peaks (harmonic components) in comparison to all other beams steered in the MVDR response, i.e., along the principal look-directions. In addition, the difference between the tallest peak of a neighboring conventional beamforming beam and the steered direction of the trip line is $< \#^\circ$ with respect to the trip line. System 10 determines whether the spectral component satisfies these conditions by comparing the bearing of the single loudest spectral component (output A 435) and the components within a bearing tolerance of each look direction (block 450) (decision block 455). If the trip line beam with the most harmonics is the tallest peak within a bearing tolerance of the trip line direction (decision block 455) system 10 proceeds to

decision block 460. Otherwise, system 10 returns to block 405 to repeat method 400.

[0077] The conditions of decision block 455 should be satisfied consecutively for at least two frames for a count to be registered, reducing chances for false target counts. This is verified by determining whether the trip line event counter was previously set, and if so, if adequate time has elapsed since the last time the counter was set. At decision block 460, system 10 determines if the trip line event counter was previously set. If not, system 10 sets the trip line event counter, records the event (block 465) and returns to block 405 to repeat method 400. If the trip line event counter has been previously set, system 10 checks at decision block 470 whether adequate time has elapsed since the last time the trip line counter was set. If insufficient time has elapsed, system 10 returns to block 405 to repeat method 400.

[0078] If the delay since the last time the trip line counter was set is greater than a predetermined threshold, at block 475 system 10 increments the trip line event counter, records the event, and returns to block 405 to repeat method 400. Once the trip line is actuated, the bearing difference between the trip line direction and bearing estimate to the tallest peak should exceed a bearing of $\#^\circ$ for the target to be counted. Otherwise, if the beam with the greatest number of harmonics becomes some other than the trip line beam, the counter is incremented by 1. If the trip line is not tripped by some predetermined threshold of frames after either one of the preceding conditions are satisfied, the target is counted and the counter is set.

[0079] A method 500 of operation of the conventional beamformer module 420 is

illustrated by the process flow chart of FIG. 4. The input to the conventional beamformer module 420 is the power spectrum for each channel of the microphone array 20 determined at block 415. System 10 generates steering vectors in the principal look directions for conventional beamforming

$$\varphi_k = \frac{\pi}{4} k, \quad \text{for } k \in \{0, K, 7\}, \text{ and steers the beams in those}$$

principal look directions at block 505.

[0080] System 10 computes output power of the conventional beamformer module 420 in the principal azimuthal directions at block 510. For example, system 10 may compute the STFT for five channels of microphone data sampled at 1024 Hz. System 10 then, in this example, computes a 2048-point SFTF using 2 seconds of data multiplied by a 2048 point Hamming Window. The STFT data is multiplied by weight vectors (Equation 8) to obtain the output power of the conventional beamformer module 420 (Equation 9), block 510.

[0081] At block 515, system 10 computes the background noise level. System 10 computes CFAR noise floor to find spectral peaks that are #dB above the threshold level. At block 520, system 10 computes the signal to noise ration (SNR) for each spectral component. System 10 then retains components greater than a predetermined threshold at block 525, and retains bearings to the single loudest spectral component at block 530. The bearing to the single loudest spectral component determined at block 530 is the output A 435.

[0082] At block 535, system 10 uses the computational output of block 510 to

compute the bearings to all spectral components. System 10 computes bearing estimates to peaks in a range, for example, of 0.5 Hz and 249.5 Hz based on the narrowband power distribution within the 8 beams (determined at block 510). The bearings to all spectral components determined at block 535 are the output B 440 of the conventional beamformer module 420 .

[0083] A method 600 of operation of the adaptive beamformer module 425 is illustrated by the process flow chart of FIG. 4. The input to the adaptive beamformer module 425 is the power spectrum for each channel of the microphone array 20 determined at block 415. The STFT frequency data calculated at block 410 is used to compute a local noise floor for each microphone element in the microphone array 20 using a rank-order filter (block 605). System 10 then generates steering vectors in the principal azimuthal directions and a "trip line" direction at block 610. At block 615, system 20 computes a correlation matrix with regularization at each frequency using spectral components across all frequencies. System 10 computes the optimum weight vectors at block 620. Using previously computed array steering vectors, system 10 steers the beams and computes the output power of the adaptive beamformer module 425 in the principal look directions and trip line direction (block 625). Background noise level or noise floor is estimated at block 630 using the same Constant False Alarm Rate (CFAR) threshold constraint as the conventional beamformer module 420. System 10 then computes the signal to noise ration for each spectral component at block 635.

[0084] As an example, consider responses from a conventional beamformer (solid line) and adaptive beamformer (dashed line) implemented for system 10 and

obtained at a single snapshot in time in FIGS. 6A, 6B, 6C, 6D, 6E, 6F, 6G, 6H using previously collected signature data. At the time that the data is extracted, three T72s tanks are in close proximity to a sensor at ranges of 104 m, 50 m, and 121 m. In addition, two ZILs have passed the sensor of interest and are located at about 700 m from the array. The T-72 at closest point of approach is at a bearing of 3.7° relative to the array, while the louder of the two remaining T-72s (due to the location of the exhaust) is located at a bearing angle of 74° .

[0085] FIGS. 6A, 6B, and 6C show that system 10 maintains a more directed response when compared to the conventional beamformer system at frequencies corresponding to harmonics propagating from the look direction (i.e. 0 degrees). As shown in FIGS. 6D, 6E, and 6F, system 10 places deep nulls in the direction of harmonics from the louder T-72 at 74° at $f = 109.5$ Hz and $f = 146.5$ Hz. In addition, system 10 places deep nulls in the direction of a harmonic component $f = 134$ Hz corresponding to the 2 ZILs that are at a bearing of 99° relative to the array. FIGS. 6G and 6H illustrate the difference of the system 10 response at two very closely spaced frequencies: $f = 146.5$ corresponding to an interfering T72 at 74° and $f = 141$ Hz corresponding to a harmonic component from the T-72 in interest. System 10 allows the acoustic sensors to obtain coverage narrow enough to pass harmonics corresponding to the target of interest and monitor the passage of the individual signal components into and out of the "trip line" steered to a fixed line of bearing angle.

[0086] It is to be understood that the specific embodiments of the invention that have been described are merely illustrative of certain application of the

principle of the present invention. Numerous modifications may be made to the system and method for detecting extended range motion and counting moving objects using an acoustics microphone array invention described herein without departing from the spirit and scope of the present invention.